

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/532,944  
Source: IFWP  
Date Processed by STIC: 7/6/06

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 07/06/2006

PATENT APPLICATION: US/10/532,944

TIME: 10:16:32

Input Set : F:\56446-20039.00 - Seqlist.txt

Output Set: N:\CRF4\07062006\J532944.raw

```

4 <110> APPLICANT: GRAY, Kevin A.
5     ABOUSHADI, Nahla
6     GARRETT, James B.
8 <120> TITLE OF INVENTION: AMYLASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
9     FOR MAKING AND USING THEM
11 <130> FILE REFERENCE: 564462003900
13 <140> CURRENT APPLICATION NUMBER: US 10/532,944
C--> 14 <141> CURRENT FILING DATE: 2005-04-26
16 <150> PRIOR APPLICATION NUMBER: PCT/US03/33150
17 <151> PRIOR FILING DATE: 2003-10-15
19 <150> PRIOR APPLICATION NUMBER: US 60/423,626
20 <151> PRIOR FILING DATE: 2002-10-31
22 <160> NUMBER OF SEQ ID NOS: 24
24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1710
28 <212> TYPE: DNA
29 <213> ORGANISM: Bacteria
31 <400> SEQUENCE: 1
32 atgcagcgcc atcgcagaag gtgcagggct aagcttgctg ggttcgtttt ggcaccccgt      60
33 ttggcgggtg catggaagcc cggagggggg ccctcgatgt cgcagactcc atggtggcgc      120
34 ggtgccgtca tctaccagat ctaccgcgc agtttcctcg acgccaacgg cgacggggtc      180
35 ggtgacctgc ccggcatcat cgaccggctg gagtacgtgg ccgcgctggg cgtggacgcc      240
36 atctgggtct cgccgttctt cacctcgccg atggccgatt tcggctacga catcgccgac      300
37 catcgcgacg tggaccgcgt gtttggcacg ctggccgatt tcgaccggct gctggccaag      360
38 gcgcacgcgc tgggcctgaa ggtgatgac caccaggtgt tcagccacac ctcgatcgac      420
39 cagcgcctgt tccgtgagag ccggcaggac gcaccaaatc cgaaggcgga ctggtacgtg      480
40 tgggcccgacc cgcgcgagga cggcacgccg cccaacaact ggatgtcgat cttcggcggg      540
41 gtggcctggc aatgggagcc gcgccgggag cagtacttcc tgcacaactt cctggccgac      600
42 cagccggacc tggatttcca caaccggcg gtgcagcagg ccacgctgga ctacgtgcgc      660
43 ttctggctgg accggggcgt ggacgggttc cgcctggact cgatcaactt ctgcttccac      720
44 gacgcgcagt tgcgcgacaa cccggccaag ccgctggaaa agcgcgtcgg ccgtggcttc      780
45 agcgcggaca atccgtacgc ctaccagtac cactactaca acaacaccca gccggagAAC      840
46 atcggcttca tcgagcgccct gcgtgggttg ctggacgaat acccgggcac cgtgagcctg      900
47 ggcgagatct cggccgagga ctcgctggcc accaccgccg agtacaccgc gccggggccgc      960
48 ctgcacatgg gctacagctt cgagctgctg gtgaaggatt tcagcgccgg ctacatccgc     1020
49 gacaccgtgt cgcggctgga agcgacgatg accgaaggct ggccgtgctg ggcgatctcc     1080
50 aaccacgacg tggagcgtgc ggtcactcgc tggggcggcc atccggcccgc gcccgggctg     1140
51 gcgcggatgc tgggtggcgt gctgtgctcg ctgctggct cgatctgcct gtaccagggc     1200
52 gaggagctgg gcctgggcga ggcggacgtg ccgttcgagg cgctgcagga cccgtatggc     1260
53 atcaccttct ggccgaactt caagggccgc gacggctgcc gcacgccgat gccgtggatc     1320
54 gatgcgccgt tggcaggggt caccagcggt gagccgtggc tgccgattcc ggccgagcac     1380
55 cgcgcgcgcg cgggtggcggg gcaggagcac gaccgcact cggtgttgaa cgcgttccgc     1440

```

## RAW SEQUENCE LISTING

DATE: 07/06/2006

PATENT APPLICATION: US/10/532,944

TIME: 10:16:32

Input Set : F:\56446-20039.00 - Seqlist.txt

Output Set: N:\CRF4\07062006\J532944.raw

```

56 cagttcctgg catggcgag gacgatgccg acgctgctgg tgggagacat cgtcttcctg 1500
57 cagacggccg agccgggtgct gatgttcgag cgccggcatg cgggggagac gctgctgctg 1560
58 gccttcaacc tggcgccga caccgcgcgc gtggcgctgc ccgccggcag ctggcagccg 1620
59 atgcacgtgc cgggcccga cgtgggcccag gccgacggcg ggacgttggt actgccggcg 1680
60 cagtcgatgt actgcgcgcg cctgggctga 1710
62 <210> SEQ ID NO: 2
63 <211> LENGTH: 569
64 <212> TYPE: PRT
65 <213> ORGANISM: Bacteria
67 <220> FEATURE:
68 <221> NAME/KEY: SIGNAL
69 <222> LOCATION: (1)...(24)
71 <400> SEQUENCE: 2
72 Met Gln Arg His Arg Arg Arg Cys Arg Ala Lys Leu Val Gly Phe Val
73 1 5 10 15
74 Leu Ala Pro Arg Leu Ala Gly Ala Trp Lys Pro Gly Gly Gly Pro Ser
75 20 25 30
76 Met Ser Gln Thr Pro Trp Trp Arg Gly Ala Val Ile Tyr Gln Ile Tyr
77 35 40 45
78 Pro Arg Ser Phe Leu Asp Ala Asn Gly Asp Gly Val Gly Asp Leu Pro
79 50 55 60
80 Gly Ile Ile Asp Arg Leu Glu Tyr Val Ala Ala Leu Gly Val Asp Ala
81 65 70 75 80
82 Ile Trp Val Ser Pro Phe Phe Thr Ser Pro Met Ala Asp Phe Gly Tyr
83 85 90 95
84 Asp Ile Ala Asp His Arg Asp Val Asp Pro Leu Phe Gly Thr Leu Ala
85 100 105 110
86 Asp Phe Asp Arg Leu Leu Ala Lys Ala His Ala Leu Gly Leu Lys Val
87 115 120 125
88 Met Ile Asp Gln Val Phe Ser His Thr Ser Ile Asp His Ala Trp Phe
89 130 135 140
90 Arg Glu Ser Arg Gln Asp Arg Thr Asn Pro Lys Ala Asp Trp Tyr Val
91 145 150 155 160
92 Trp Ala Asp Pro Arg Glu Asp Gly Thr Pro Pro Asn Asn Trp Met Ser
93 165 170 175
94 Ile Phe Gly Gly Val Ala Trp Gln Trp Glu Pro Arg Arg Glu Gln Tyr
95 180 185 190
96 Phe Leu His Asn Phe Leu Ala Asp Gln Pro Asp Leu Asp Phe His Asn
97 195 200 205
98 Pro Ala Val Gln Gln Ala Thr Leu Asp Tyr Val Arg Phe Trp Leu Asp
99 210 215 220
100 Arg Gly Val Asp Gly Phe Arg Leu Asp Ser Ile Asn Phe Cys Phe His
101 225 230 235 240
102 Asp Ala Gln Leu Arg Asp Asn Pro Ala Lys Pro Leu Glu Lys Arg Val
103 245 250 255
104 Gly Arg Gly Phe Ser Ala Asp Asn Pro Tyr Ala Tyr Gln Tyr His Tyr
105 260 265 270
106 Tyr Asn Asn Thr Gln Pro Glu Asn Ile Gly Phe Ile Glu Arg Leu Arg
107 275 280 285

```

## RAW SEQUENCE LISTING

DATE: 07/06/2006

PATENT APPLICATION: US/10/532,944

TIME: 10:16:32

Input Set : F:\56446-20039.00 - Seqlist.txt

Output Set: N:\CRF4\07062006\J532944.raw

```

108 Gly Leu Leu Asp Glu Tyr Pro Gly Thr Val Ser Leu Gly Glu Ile Ser
109      290                      295                      300
110 Ala Glu Asp Ser Leu Ala Thr Thr Ala Glu Tyr Thr Ala Pro Gly Arg
111 305                      310                      315                      320
112 Leu His Met Gly Tyr Ser Phe Glu Leu Leu Val Lys Asp Phe Ser Ala
113                      325                      330                      335
114 Gly Tyr Ile Arg Asp Thr Val Ser Arg Leu Glu Ala Thr Met Thr Glu
115                      340                      345                      350
116 Gly Trp Pro Cys Trp Ala Ile Ser Asn His Asp Val Glu Arg Ala Val
117      355                      360                      365
118 Thr Arg Trp Gly Gly His Pro Ala Arg Pro Arg Leu Ala Arg Met Leu
119      370                      375                      380
120 Val Ala Leu Leu Cys Ser Leu Arg Gly Ser Ile Cys Leu Tyr Gln Gly
121 385                      390                      395                      400
122 Glu Glu Leu Gly Leu Gly Glu Ala Asp Val Pro Phe Glu Ala Leu Gln
123                      405                      410                      415
124 Asp Pro Tyr Gly Ile Thr Phe Trp Pro Asn Phe Lys Gly Arg Asp Gly
125      420                      425                      430
126 Cys Arg Thr Pro Met Pro Trp Ile Asp Ala Pro Leu Ala Gly Phe Thr
127      435                      440                      445
128 Ser Gly Glu Pro Trp Leu Pro Ile Pro Ala Glu His Arg Ala Ala Ala
129      450                      455                      460
130 Val Ala Val Gln Glu His Asp Pro His Ser Val Leu Asn Ala Phe Arg
131 465                      470                      475                      480
132 Gln Phe Leu Ala Trp Arg Arg Thr Met Pro Thr Leu Leu Val Gly Asp
133                      485                      490                      495
134 Ile Val Phe Leu Gln Thr Ala Glu Pro Val Leu Met Phe Glu Arg Arg
135      500                      505                      510
136 His Ala Gly Glu Thr Leu Leu Leu Ala Phe Asn Leu Ala Ala Asp Thr
137      515                      520                      525
138 Ala Arg Val Ala Leu Pro Ala Gly Ser Trp Gln Pro Met His Val Pro
139      530                      535                      540
140 Gly Pro Asp Val Gly Gln Ala Asp Gly Gly Thr Leu Val Leu Pro Ala
141 545                      550                      555                      560
142 Gln Ser Met Tyr Cys Ala Arg Leu Gly
143      565
145 <210> SEQ ID NO: 3
146 <211> LENGTH: 1293
147 <212> TYPE: DNA
148 <213> ORGANISM: Unknown
150 <220> FEATURE:
151 <223> OTHER INFORMATION: Obtained from an environmental sample
153 <400> SEQUENCE: 3
154 atgagtctgt ggcgtgcgct cgtcgcgttt gttctgctgg ccgtcgccat gcccgcatTT      60
155 gccgacgtgg tcgcgaccgc ctcttcgcca ggcgatgtcc tcaaggTgga gatcaccacc      120
156 aacggcgagg gccgcacTcg ctatgcggTc acccggtcTg gcaagccggT aatcggcgag      180
157 agccacTctg gattctctct ggccgacgcg ccgcagctgc tgcgcaactt ccaggTcgtc      240
158 gatcaggcca cccggacTtt cgacgaaacg tgggagcagc cgtggggggga gtggcgcacg      300
159 gtccgcaacc actacaacga gctcgcgatc accttcgagg agaagaccaa gctccatcgg      360

```

## RAW SEQUENCE LISTING

DATE: 07/06/2006

PATENT APPLICATION: US/10/532,944

TIME: 10:16:32

Input Set : F:\56446-20039.00 - Seqlist.txt

Output Set: N:\CRF4\07062006\J532944.raw

```

160 cggatgcgga tcgttttttcg cctgttcgat gaagggatcg gcttttcgcta cgagcttccc 420
161 cggcagccga acctggcgca cgccaacatc gccgaggaac tgacctcagtt caacgtcgcg 480
162 cgaccgggca cggcctggtg ggcaccggcg ttcgaatcca acccgagga atatctctac 540
163 aaccagaccc cgatcgacgg tgtcgcgatt gcgatgactc cattcacgat gcggttcgag 600
164 gacgggactc acctcagcat ccacgaagcc gcgctggtcg actattccgg gatgaacgtc 660
165 acgcgtgtcc agggcacgaa cttcaaggcc atcctgacgc ccggttcgat gggcccaaaa 720
166 gtctcccgcg ataccccatt cgagaccccg tggcggggtca tcctgatcag ccccgacgct 780
167 gcgcacctct acgaatcgaa caggctgata ctcaacctca acgaacccaa caagctcggc 840
168 gacgtcagct ggggtccaccc gcgcaaatat gtcggcatct ggtggggcat gcacctcgat 900
169 acccagagtt gggcctcggg gccgaagcac ggcgcgacca ccgcttatgc gaagcgaatg 960
170 atcgatttcg cggcgacgaa cggttttacc gggctgctcg tcgaaggctg gaacaagggg 1020
171 tgggacggag actggttcgc gaccggcgac gatttcagct tcaccgaacc ctatcccgat 1080
172 ttcgacatcc gagccgtcgc ggctacacg ctcaggaaag gcgtccacct catcgggcat 1140
173 cacgaaacca gcggcaacat cgcccactac gagcaacagc tcgatgcggc gctcgatctc 1200
174 gaccgccagc tcggcatcga caggtgaag acgggctacg tctcgatgc cggcggcac 1260
175 caggcgctcg ggcccagcg caggatccaa aga 1293
177 <210> SEQ ID NO: 4
178 <211> LENGTH: 431
179 <212> TYPE: PRT
180 <213> ORGANISM: Unknown
182 <220> FEATURE:
183 <223> OTHER INFORMATION: Obtained from an environmental sample
185 <400> SEQUENCE: 4
186 Met Ser Leu Trp Arg Ala Leu Val Ala Phe Val Leu Leu Ala Val Ala
187 1 5 10 15
188 Met Pro Ala Phe Ala Asp Val Val Ala Thr Ala Ser Ser Pro Gly Asp
189 20 25 30
190 Val Leu Lys Val Glu Ile Thr Thr Asn Gly Glu Gly Arg Ile Gly Tyr
191 35 40 45
192 Ala Val Thr Arg Leu Gly Lys Pro Val Ile Gly Glu Ser His Leu Gly
193 50 55 60
194 Phe Leu Leu Ala Asp Ala Pro Gln Leu Leu Arg Asn Phe Gln Val Val
195 65 70 75 80
196 Asp Gln Ala Thr Arg Thr Phe Asp Glu Thr Trp Glu Gln Pro Trp Gly
197 85 90 95
198 Glu Trp Arg Thr Val Arg Asn His Tyr Asn Glu Leu Ala Ile Thr Phe
199 100 105 110
200 Glu Glu Lys Thr Lys Leu His Arg Arg Met Arg Ile Val Phe Arg Leu
201 115 120 125
202 Phe Asp Glu Gly Ile Gly Phe Arg Tyr Glu Leu Pro Arg Gln Pro Asn
203 130 135 140
204 Leu Ala His Ala Asn Ile Ala Glu Glu Leu Thr Gln Phe Asn Val Ala
205 145 150 155 160
206 Arg Pro Gly Thr Ala Trp Trp Ala Pro Ala Phe Glu Ser Asn Arg Glu
207 165 170 175
208 Glu Tyr Leu Tyr Asn Gln Thr Pro Ile Asp Gly Val Ala Ile Ala Met
209 180 185 190
210 Thr Pro Phe Thr Met Arg Phe Glu Asp Gly Thr His Leu Ser Ile His
211 195 200 205

```

## RAW SEQUENCE LISTING

DATE: 07/06/2006

PATENT APPLICATION: US/10/532,944

TIME: 10:16:32

Input Set : F:\56446-20039.00 - Seqlist.txt

Output Set: N:\CRF4\07062006\J532944.raw

212 Glu Ala Ala Leu Val Asp Tyr Ser Gly Met Asn Val Thr Arg Val Gln  
 213       210                               215                               220  
 214 Gly Thr Asn Phe Lys Ala Ile Leu Thr Pro Gly Ser Met Gly Pro Lys  
 215 225                               230                               235                               240  
 216 Val Ser Arg Asp Thr Pro Phe Glu Thr Pro Trp Arg Val Ile Leu Ile  
 217                               245                               250                               255  
 218 Ser Pro Asp Ala Ala His Leu Tyr Glu Ser Asn Arg Leu Ile Leu Asn  
 219                               260                               265                               270  
 220 Leu Asn Glu Pro Asn Lys Leu Gly Asp Val Ser Trp Val His Pro Arg  
 221                               275                               280                               285  
 222 Lys Tyr Val Gly Ile Trp Trp Gly Met His Leu Asp Thr Gln Ser Trp  
 223       290                               295                               300  
 224 Ala Ser Gly Pro Lys His Gly Ala Thr Thr Ala Tyr Ala Lys Arg Met  
 225 305                               310                               315                               320  
 226 Ile Asp Phe Ala Ala Thr Asn Gly Phe Thr Gly Leu Leu Val Glu Gly  
 227                               325                               330                               335  
 228 Trp Asn Lys Gly Trp Asp Gly Asp Trp Phe Ala Thr Gly Asp Asp Phe  
 229                               340                               345                               350  
 230 Ser Phe Thr Glu Pro Tyr Pro Asp Phe Asp Ile Arg Ala Val Ala Ala  
 231                               355                               360                               365  
 232 Tyr Ser Leu Arg Lys Gly Val His Leu Ile Gly His His Glu Thr Ser  
 233       370                               375                               380  
 234 Gly Asn Ile Ala His Tyr Glu Gln Gln Leu Asp Ala Ala Leu Asp Leu  
 235 385                               390                               395                               400  
 236 Asp Arg Gln Leu Gly Ile Asp Thr Val Lys Thr Gly Tyr Val Ser Asp  
 237                               405                               410                               415  
 238 Ala Gly Gly Ile Gln Ala Leu Gly Pro Asp Gly Arg Ile Gln Arg  
 239                               420                               425                               430

241 &lt;210&gt; SEQ ID NO: 5

242 &lt;211&gt; LENGTH: 1839

243 &lt;212&gt; TYPE: DNA

244 &lt;213&gt; ORGANISM: Bacteria

246 &lt;400&gt; SEQUENCE: 5

247 atgcatatcc gcatcgccag aatgctgaca gcgctgtaca tcatacgcgg ccgaaaggag 60  
 248 ttgcatatga cctcaacaa caccatgcc gactggtgga aacaggcggg ggtctaccag 120  
 249 gtctacccgc gcagcttcaa ggatgtgaac ggtgacggtc tgggcgacat cgccggcgtg 180  
 250 accgaccgca tcccctatct caaggagctg ggcgtcgacg cgatctggtt gtcgccgttc 240  
 251 taccctgccg agctggcgga cggcggtac gacgtcatcg actaccgca cgtggaccgc 300  
 252 cgctggggct ccatggacga cttcgacgcc atggccgcag ccgcgcacga ggccggcatg 360  
 253 aagtgatcg tggacatcgt gccgaaccac acctccgacc ggcacgtctg gtttgaggag 420  
 254 gcgctggcgg cagaaccggg ctcccccgcg cgcgaccggt acatcttccg cgacggtctg 480  
 255 gggaacacg gcgagctgcc cccgaacgac tggcagtcga tcttcggcgg cggcgcacatg 540  
 256 gaacgggtgc cggacggcca gtggtacctg cacatgttcg ccaaggaaca gcccgacctc 600  
 257 aactggaaga acccgagggt ccacgaggag ttcaagaaga ccctgcgttt ctggtcggac 660  
 258 catggcgcgg acgggttccg catcgacgtg gcgcacggcc tggccaagga cctcgacagc 720  
 259 gcgccgctga ccgaactggc gttgaagggc aaccgcgtgg agggactgtg ccatgacggc 780  
 260 acgaatccgc tgtgggaccg ccccgagggt cagcacatct accgcgaatg gcgcgaggtg 840  
 261 ttcaacgagt acgatccgcc gcgcttcgcc gtcggcgagg cctgggtggt gcccgaaac 900  
 262 cagcacctgt acgcctccga ggaggagctg gggcaggtgt tcaacttcga attcgccaag 960

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/532,944

DATE: 07/06/2006

TIME: 10:16:33

Input Set : F:\56446-20039.00 - Seqlist.txt

Output Set: N:\CRF4\07062006\J532944.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date